Session 3: Dealing with Reverse Causality

Ben Shepherd

Principal, Developing Trade Consultants Ltd.

ARTNeT Capacity Building Workshop for Trade Research:
Gravity Modeling

Thursday, August 26, 2010

Outline

Outline

Overview

- One of the most important OLS assumptions is that the errors are uncorrelated with the dependent variables.
- A number of possible reasons:
  - Measurement error
  - Omitted variables
  - Endogeneity ("reverse causality")
- Endogeneity is particularly important in policy settings. Pervasive in trade facilitation settings.

Overview

- Tariffs are an obvious example of an endogeneity problem in many gravity models.
- Intuition suggests that tariffs should impact negatively on bilateral trade.
- But political economy suggests that greater import penetration is likely to lead to increased demand for protection by industry lobbies (e.g., Grossman and Helpman, “protection for sale”)
- What are the consequences for estimation and inference? How can we deal with the problem?

Outline

Conceptualizing Endogeneity

- A simple example of endogeneity is the following:

$$\mathbf{A} = \mathbf{XB} + \mathbf{E}$$

$$\mathbf{X} = \mathbf{AC} + \mathbf{W}$$

- Since the current value of X depends on the current value of A, X must be influenced by current shocks to A:

$$\mathbf{X} = (\mathbf{XB} + \mathbf{E})\mathbf{C} + \mathbf{W}$$

- Thus, X and E are correlated. This violates the OLS assumptions.

Consequences of Endogeneity

- Endogeneity bias is not a “simple” violation to deal with. There is no equivalent to the robust estimators used to conduct inference in the presence of general patterns of heteroskedasticity.
- Moreover, it has serious consequences for our estimates.
- In the presence of endogeneity, OLS can produce biased and inconsistent parameter estimates. Hypotheses tests can be seriously misleading.
- All it takes is one endogenous variable to seriously distort ALL OLS estimates of a model.

Outline

Introduction

- We need some way of separating out genuinely exogenous variation in independent variables that might be endogenous.
- Ideally, we would also like some way of testing the extent to which endogeneity is a problem in our data, and ensuring that the solution we have chosen is a good one.
- Two groups of solutions:
  - Ad hoc approaches
  - Instrumental variables estimation

Ad Hoc Solutions

- If a dependent variable is potentially endogenous, it is intuitively appealing to look for a proxy that does not suffer from the same problem.
- The most common approach is to lag the suspect variables by one or more periods.
- The argument is that although current values of (eg) tariffs might be endogenous to import penetration, it is unlikely that past values of tariffs are subject to the same problem.

Ad Hoc Solutions

Advantages of Lags/Proxies

- Very simple to implement.
- Additional data requirements limited.
- Intuitively appealing.

Limitations of Lags/Proxies

- Interpretation becomes a little more difficult since the variable in the regression is only a proxy for the variable we are interested in.
- Loss of precision in some cases.
- No way of gauging empirically how serious the endogeneity problem is, and whether the solution is adequate to deal with it.

Outline

Introduction

- The best way to deal with endogeneity concerns is through instrumental variables (IV) techniques.
- The most common IV estimator is Two Stage Least Squares (TSLS).
- Intuitively, IV estimation works as follows:
  - Find a genuinely exogenous variable (instrument) that is strongly correlated with the potentially endogenous regressor.
  - Ensure that the instrument only influences the dependant variable through the potentially endogenous independent variable.

Introduction

- IV estimation is intuitively appealing, and relatively simple to implement on a technical level.
- The most difficult part lies in the selection of appropriate instruments.
- The main advantages of IV estimation are:
  - Rigor and transparency
  - Amenability to empirical testing (appropriateness of instruments, extent of endogeneity, etc.)

Two Stage Least Squares

- As the name suggests, this estimator proceeds in two steps. Both of them involve OLS.

1. Regress the potentially endogenous variables in the gravity model on the exogenous variables from the model AND a number of genuinely exogenous "instruments" equal to the number of potentially endogenous variables.
2. Run the gravity model using the predicted values from the first stage regressions in place of the potentially endogenous variables.

Two Stage Least Squares

- Why does TSLS work? Intuitively:
  - The first stage "cleanses" the endogeneity from the variables we are worried about. By using predicted values based on genuinely exogenous variables only, we obtain the exogenous part of their variation.
  - The second stage uses a variable that is now exogenous thanks to the first stage, and so the bias disappears.
- Interpretation of parameters and hypothesis testing can all take place as usual, following the same procedures as OLS.

Two Stage Least Squares

- The hard part about TSLS is not the estimation, but the data.
- For one potentially endogenous variable, we need at least one instrumental variable. For two, we need at least two, etc.
- What is a valid instrument?
  - It is strongly correlated with the potentially endogenous variable.
  - It is exogenous to the gravity model, and does not influence bilateral trade EXCEPT through the potentially endogenous variable.

Two Stage Least Squares

In practice:

1. IV estimates are only as good as the instruments used. We need to run some statistical tests and make sure the model is working properly.
2. Although it's called "two stage" least squares, avoid running the two stages separately. You will get incorrect standard errors (too small), and you might mistakenly exclude exogenous variables from the main model–a common error.
3. Use STATA's "canned" routes: ivreg, and ivreg2 (includes the tests we are about to discuss). But make sure you know what they're telling you!

Estimation in Stata

- To make life easy, use the external ivreg2 and xtivreg2 commands. Use the "first" option to get the first stage regression results.
- For a gravity model in which tariffs are potentially endogenous, we specify the following in Stata:
  - ivreg2 ln_trade ln_dist (ln_tariff = instrument) [etc.] dummies, robust cluster(dist) first
  - xtivreg2 ln_trade ln_dist (ln_tariff = instrument) [etc.], robust first

Testing TSLS Models

Overview

TSLS provides a number of useful tests that can help in deciding whether IV estimation is necessary, and whether the instruments chosen are valid.

- Endogeneity test: Is there evidence that correlation between the potentially endogenous variables and the error term is strong enough to result in substantively biased estimates?
- Instrument relevance test: Are the instruments sufficiently strongly correlated with the potentially endogenous variables?
- Exogeneity/excludability of instruments: Are the instruments genuinely uncorrelated with the main equation residuals?

Testing TSLS Models

Testing for Instrument Relevance

- TSLS is only as good as the instruments.
- In the presence of weak instruments, the TSLS estimator can actually produce worse results than simple OLS.
- So the first step in testing must be to ensure that the instruments are strongly enough correlated with the potentially endogenous variables.

Testing TSLS Models

Testing for Instrument Relevance

- To test for instrument relevance, make sure to run the first stage regressions of the potentially endogenous variables on all of the exogenous variables.
- Are the instruments individually statistically significant? Are their signs and magnitudes sensible?
- Are the instruments jointly statistically significant? Look for a high F-statistic.
- Extension: Explore test statistics designed specifically to deal with weak instruments, e.g. Stock/Yogo.

Testing TSLS Models

Testing for Endogeneity

- In gravity models, endogeneity is nearly always a problem for policies, at least in principle.
- Estimating a TSLS model implies costs:
  - Loss of precision, due to two stage estimation;
  - Possibility of incorrect inference when instruments are weak.
- It would therefore be nice to have a test that allows us to see how serious the endogeneity problem really is. Do the benefits of TSLS outweigh the costs in a particular instance?

Testing TSLS Models

Testing for Endogeneity

- We can test for the endogeneity of suspect independent variables using a Hausman test.
- H0: the regressor is exogenous; H1: it is endogenous.
- Thus, a higher value of the test statistic indicates a more serious endogeneity problem.
- Start to worry if it rejects the null at the 10% level.

Testing TSLS Models

Testing for Endogeneity

- The mechanics of the test are cumbersome.
- Intuitively, it is equivalent (in large samples) to the following procedure:
  - Run the first stage regressions and save the residuals;
  - Include the first stage residuals as additional regressors in the main equation;
  - Test the joint significance of the first stage residuals.
- In other words, the test looks at whether or not there is a correlation between the dependent variable, and that part of the suspect variables' variation that is not explained by genuinely exogenous factors.
- In Stata: specify the endog(varname) option with ivreg2 or xtivreg2.

Testing TSLS Models

Testing for Exogeneity and Excludability of Instruments

- In addition to being relevant, instruments must also be:
  - Exogenous with respect to the dependent variable in the main equation; and
  - Excludable from the main regression.
- We can use another version of the Hausman test (in fact, Hansen's J) to test whether or not these conditions jointly hold.
- The test can ONLY be run if the model is overidentified, i.e. if we have more excluded instruments than potentially endogenous variables.
- Ivreg2 and xtivreg2 automatically run this test if there are enough instruments.

Testing TSLS Models

Testing for Exogeneity and Excludability of Instruments

- H0 for Hansen's J: the instruments are uncorrelated with the main equation errors; H1: the instruments are correlated with the main equation errors.
- Thus, a high value of the test statistic tends to indicate that there is a problem with the instruments:
  - Endogeneity: They are correlated with the main equation errors because there is feedback running from the dependent variable to the instruments; and/or
  - Non-excludability: The instruments should appear in the main regression, and the test is effectively picking up an omitted variables problem.
- Be scared if Hansen's J rejects the null at the 10% level.

Testing TSLS Models

- Always test for instrument relevance first. If instruments are weak, tests like Hausman and Hansen's J can give misleading results.
- Use the Hausman test to assess the extent to which endogeneity is really a problem.
- If at all possible, ensure the model is overidentified, and test exogeneity and excludability via Hansen's J.
- If the model passes all of these tests, then it should provide a reasonable guide for causal inference.

Outline

Examples of IV Gravity Models

"Trading on Time" (Djankov et al., 2009)

- The main independent variable of interest is the Doing Business measure of trading time.
- But time is potentially endogenous to trade: an unexpected increase in trade volumes can lead to congestion, and longer wait times.
- Solution: Use the number of administrative formalities (documents) as an instrument for time.
  - Time and documents are very strongly correlated.
  - Test results suggest that exogeneity/excludability is also satisfied. Intuitively, though, excludability is more questionable: wouldn't very complex and burdensome documentation requirements constitute an independent source of trade costs?

Examples of IV Gravity Models

"Trading on Time" (Djankov et al., 2009)

Independent Variables	*Unrestricted Sample*		
	(1)	(2)	(3)
Ratio_Time		**-0.414** **	**-0.512****
		(-6.37)	**(-5.34)**
Ratio_GDP	1.110 **	1.140 **	1.147 **
	(36.88)	(38.01)	(37.48)
Ratio_GDPC	0.263 **	0.106 *	0.069
	(4.88)	(1.70)	(1.02)
Ratio_Distance	-1.232 **	-1.233 **	-1.233 **
	(-21.47)	(-21.15)	(-21.01)
Contiguity	0.489 **	0.473 **	0.469**
	(5.96)	(5.80)	(5.76)
Language	0.663 **	0.715 **	0.727**
	(8.10)	(8.54)	(8.58)
Colony	0.487 **	0.544 **	0.557 **
	(5.23)	(6.01)	(6.16)
Landlocked	-0.429 **	-0.368 **	-0.353 **
	(-4.62)	(-4.10)	(-3.85)
Ratio_Remoteness	-0.454 **	-0.334 **	-0.306 **
	(-4.76)	(-3.73)	(-3.26)
Instruments			
No of Required Signatures for Exports and Imports to Take Place			Yes
Overidentification Test			*0.098*
P-Value			*0.75408*
R^2 in the First Stage			0.5243
R^2	0.5053	0.5114	0.5110
No of Obs.	42,675	42,675	42,675

Examples of IV Gravity Models

"Trading on Time" (Djankov et al., 2009)

- In the latest version of their paper, they take a different approach.
- They instrument for trading times using times in neighboring countries.
- Their results remain robust to IV regression.

Examples of IV Gravity Models

"Trading on Time" (Djankov et al., 2009)

Independent Variables	*Regional Agreement Sample*		*Regional Agreement and Income Group*	*Landlocked Country Sample*		
	(1)	(2)	(3)	(4)	(5)	(6)
Ratio_time		-0.484 ***	-0.412 ***	-0.559		-1.034 **
		(-7.17)	(-5.34)	(-1.43)		(-1.96)
Ratio_export time in neighbors					-1.869 **	
					(-2.32)	
Ratio_GDP	1.146 ***	1.170 ***	1.134 ***	1.818 ***	2.001 ***	1.847 ***
	(41.38)	(43.09)	(33.92)	(8.17)	(8.56)	(7.75)
Ratio_GDPC	0.315 ***	0.116 *	0.446 ***	-0.891 ***	-1.001 ***	-0.878 ***
	(5.82)	(1.81)	(2.93)	(-3.84)	(-4.02)	(-3.66)
Distance	-1.272 ***	-1.255 ***	-1.296 ***	-0.833 ***	-0.763 ***	-0.731 **
	(-23.05)	(-22.06)	(-20.53)	(-2.95)	(-2.61)	(-2.48)
Contiguity	0.533 ***	0.533 ***	0.471 ***	1.598 ***	1.643 ***	1.986 ***
	(6.41)	(6.40)	(4.59)	(3.12)	(3.22)	(3.61)
Language	0.720 ***	0.758 ***	0.670 ***	0.414	0.526 **	0.573 **
	(8.84)	(9.13)	(8.42)	(1.14)	(2.12)	(2.22)
Colony	0.503 ***	0.566 ***	0.528 ***	0.469 *	0.398	0.460
	(5.49)	(6.38)	(6.03)	(1.70)	(1.11)	(1.27)
Landlocked	-0.387 ***	-0.340 ***	-0.341 ***			
	(-4.14)	(-3.82)	(-2.83)			
Instrument: Transit time In neighbors	No	No	No	No	No	Yes

Outline

Examples of IV Gravity Models
"Contract Enforcement" (Ranjan & Lee, 2007)

- Contract enforcement can affect international trade transaction costs.
- But, institutional development might be endogenous to trade.
- So they instrument for contract enforcement using mortality figures for European settlers during the colonial era.

Examples of IV Gravity Models
"Contract Enforcement" (Ranjan & Lee, 2007)

- Why use settler mortality as an instrument for current contract enforcement?
- In a famous paper, Acemoglu Johnson and Robinson (AER, 2001) show that there is in fact a strong correlation between them, and that it is not just a statistical artefact:
  - European countries tended to operate extractive colonies in harsh (malarial) environments, and settler colonies in more temperate ones;
  - As a result, high mortality environments tended to see lesser importation of European governance institutions;
  - And so we now observe a strong correlation between settler mortality in the colonial era, and current measures of governace.

Examples of IV Gravity Models
"Contract Enforcement" (Ranjan & Lee, 2007)

We are interested in the last three columns:

	Importing country effect random			Importing country effect fixed		
	Differentiated	Reference	Organized	Differentiated	Reference	Organized
HF_i	**1.02****	**0.88****	**1.04****			
	(1.05)	(0.11)	(0.12)			
HF_j	**2.62****	**2.38****	**1.78****	**2.53****	**2.29****	**1.70****
	(3.01)	(0.26)	(0.22)	(0.04)	(0.04)	(0.04)

Examples of IV Gravity Models
"Contract Enforcement" (Ranjan & Lee, 2007)

Compare the first three columns with the last three from the previous slide

	Differentiated	Reference	Organized
HF_j	**4.42****	**4.04****	**3.26****
	(0.12)	(0.12)	(0.11)

Summary

- Endogeneity can be a very serious issue in gravity models.
- It is not just “academic”: it can result in wrong coefficient signs, and significant under- or over-estimation of policy effects.
- Basically two ways to deal with it:
  - Ad hoc solutions (proxies, lags);
  - IV estimation

Summary

- IV estimation is only useful if the instruments are good. They must be:
  - Strongly correlated with the potentially endogenous variables;
  - Genuinely exogenous to the model, and correlated with trade only through the potentially endogenous variables.
- Whenever possible, overidentify the model (i.e., include more instruments than potentially endogenous variables): it allows you to test for instrument exogeneity and excludability.
- Footnote: GMM is a cooler way of doing IV estimation. Identical to TSLS for just identified models, but can give more precise estimates with over-identified models.